Schilling, E.E., N. Mattson, and A. Floden. 2014. Barcoding the Asteraceae of Tennessee, tribe Coreopsideae. Phytoneuron 2014-101: 1–6. Published 20 October 2014. ISSN 2153 733X

BARCODING THE ASTERACEAE OF TENNESSEE, TRIBE COREOPSIDEAE

EDWARD E. SCHILLING, NICHOLAS MATTSON, AARON FLODEN
Herbarium TENN
Department of Ecology & Evolutionary Biology
University of Tennessee
Knoxville, Tennessee 37996
eschilling@utk.edu; afloden@utk.edu

ABSTRACT

Results from barcoding studies of tribe Coreopsideae for the Tennessee flora using the nuclear ribosomal ITS marker are presented and include the first complete reports for 2 of the 20 species of the tribe that occur in the state, as well as updated reports for several others. Sequence data from the ITS region separate most of the species of *Bidens* in Tennessee from one another, but species of *Coreopsis*, especially those of sect. *Coreopsis*, have ITS sequences that are identical (or nearly so) to at least one congener. Comparisons of sequence data to GenBank records are complicated by apparent inaccuracies of older sequences as well as potentially misidentified samples. Broad survey of *C. lanceolata* from across its range showed little variability, but the ITS sequence of a morphologically distinct sample from a Florida limestone glade area was distinct in lacking a length polymorphism that was present in other samples.

Tribe Coreopsideae is part of the Heliantheae alliance and earlier was often included in an expanded Heliantheae (Anderberg et al. 2007) in which it was usually treated as a subtribe (Crawford et al. 2009). The tribe shows a small burst of diversity in the southeastern USA involving *Bidens* and *Coreopsis* sect. *Coreopsis*. The current study continues the effort to characterize the levels and patterns of molecular diversity found in species of Asteraceae in Tennessee and southeastern North America (Schilling & Floden 2012, 2014; Schilling 2013) and to assess the potential of the nuclear ribosomal ITS region as a molecular barcode to identify species.

Coreopsideae in Tennessee includes three genera and about 20 species (Chester et al. 2009). The two species of *Cosmos* have been collected as garden escapes and may be adventive in disturbed sites. *Bidens* and *Coreopsis* have up to nine species each, all of which may be native, although this is not certain for some, particularly of *Coreopsis*, which are popular garden plants. The generic limits and separation of *Bidens* and *Coreopsis* continue to be problematic (Crawford et al. 2009). Two species of *Coreopsis*, *C. delphiniifolia* and *C. latifolia*, are considered to be rare both globally and in the state (Crabtree 2012).

The goal of this study was to conduct a survey of variation for the ITS marker for all species of Coreopsideae that occur in Tennessee. Particular emphasis was placed on *Coreopsis* and *Bidens*, both of which include modest radiations in southeastern North America and for which species boundaries are still under discussion. Within *Coreopsis*, a broader regional sampling of *C. lanceolata* was undertaken to help evaluate whether a morphologically distinct entity that occurs in limestone glade habitats in Florida (Johnson et al. 2013) is also distinct at the molecular level. Also included were several samples of horticultural cultivars that might be encountered as escapes.

Materials and methods

DNA was extracted from leaf samples either collected fresh or taken from herbarium specimens (Table 1). DNA extraction, PCR amplification, and sequencing protocols followed Schilling and Floden (2011). Samples that had length polymorphisms in the ITS region were sequenced with multiple primers. GenBank accession numbers are provided in Table 1. Although

this study was not designed to undertake a rigorous phylogenetic analysis, a maximum likelihood tree was generated using MEGA 6 (Tamura et al. 2013) to provide a convenient way to make a comparative visualization of the sequence results. The tree was rooted using a sample of *Gaillardia* (Helenieae), another member of the Heliantheae alliance. The analysis also utilized sequences deposited at GenBank of conspecific samples or closely related species of Coreopsideae.

Table 1. Plant material used for ITS barcoding studies of Coreopsideae. All voucher specimens at TENN.

Species	DNA#	Genbank	Voucher info
BIDENS L.			
B. aristosa Britton	2544	KM347897	*Schilling 07-DNA2544*, Knox Co., TN
B. aristosa Britton	3817	KM347898	*Floden 1707*, Campbell Co., TN
B. bipinnata L.	3818	KM347899	*Estes and Beck 6410*, Fentress Co., TN
B. bipinnata L.	3530	KM347900	*DeSelm 02-219*, Bradley Co., TN
B. cernua L.	3819	KM347901	*Floden 1415*, Campbell Co., TN
B. cernua L.	3531	KM347902	*Webb 6912*, Benton Co., TN
B. discoidea Britton	3532	KM347903	*Webb 6911,* Benton Co., TN
B. discoidea Britton	3820	KM347904	*DeSelm 02-287*, Claiborne Co., TN
B frondosa L.	3533	KM347905	*Webb 7628*, Henry Co., TN
B frondosa L.	3821	KM347906	*Estes and Beck 8402*, Grundy Co., TN
B frondosa L.	2624	KM347907	*Schilling 7-BC2624,* Knox Co., TN
B. polylepis S.F. Blake	3822	KM347908	*Souza 86-609*, Dickson Co., TN
B. polylepis S.F. Blake	2588	KM347909	*Schilling 07-DNA2588,* Knox Co., TN
B. tripartita L.	3823	KM347910	*Webb 6910*, Benton Co., TN
COREOPSIS L.			
C. auriculata L.	3016	KM347910	*Schilling 10-DNA3016*, Campbell Co., TN
C. basalis S.F. Blake	3816	KM347913	*Schilling 13-DNA3816*, cv. Sunburst
C. delphiniifolia Scherff	3091	KM347914	*Wofford and Clebsch 91-19*, Polk Co., TN
C. sp. nov. –Florida Glades	3622	KM347915	*Johnson 13,* Jackson Co., FL
C. grandiflora Hogg ex Sweet	3092	KM347916	*DeSelm s.n. 5/14/06*, Bradley Co., TN
C. grandiflora Hogg ex Sweet	3826	KM347917	*Estes and Beck 7618*, Marion Co., TN
C. grandiflora Hogg ex Sweet	3815	KM347918	*Schilling 13-DNA3815*, Greenhouse plant
C. grandiflora var. *inclinata* J.R.Allison	3679	KM347919	*Allison 12086*, Bibb Co., AL
C. lanceolata L.	3093	KM347920	*McCoy s.n. 5/30/07*, Rutherford Co., TN
C. lanceolata L.	3621	KM347921	*Johnson 13*, Gadsden Co., FL
C. lanceolata var. *villosa* Michx.	3650	KM347922	*Schmidt97-132*, Mackinac Co., MI
C. lanceolata L.	3651	KM347923	*Stutts 351*, Muscogee Co., GA
C. lanceolata L.	3652	KM347924	*Thomas 88326*, Webster Par., LA
C. lanceolata L.	3653	KM347925	*Thomas 159401*, Nevada Co., AR
C. lanceolata L.	3863	KM347926	*DeSelm s.n.2000*, Monroe Co., TN
C. lanceolata L.	3867	KM347927	*R. Kral 1983*, Warren Co., TN
C. lanceolata L.	3886	KM347928	*Bailey s.n.,* Perry Co., TN
C. lanceolata L.	3868	KM347929	*R. Kral 385*, 1975, Sumner Co., TN
C. lanceolata L. cv. Sterntaler	3813	KM347930	*Schilling 13-DNA3813*, Garden plant
C. lanceolata L. cv. Sunfire	3814	KM347931	*Schilling 13-DNA3814*, Garden plant
C. lanceolata L.	3827	KM347932	*Souza 88-164*, Dickson Co., TN
C. lanceolata var. *villosa* Michx.	3649	KM347933	*Godfrey 55430*, Calhoun Co., FL
C. lanceolata var. *villosa* Michx.	3864	KM347934	*Rhinehart s.n. 2001*, Sequatchie Co., TN
C. lanceolata var. *villosa* Michx.	3865	KM347935	*Mcneilus 00-460*, Polk Co., TN
C. lanceolata var. *villosa* Michx.	3866	KM347936	*McCoy s.n. 5/30/07*, Rutherford Co., TN
C. lanceolata var. *villosa* Michx.	3869	KM347937	*R.Kral 385, 1975,* Sumner Co., TN
C. latifolia Michx.	3094	KM347938	*Murrell* et al *948a*, Polk Co., TN
C. latifolia Michx.	3828	KM347939	*Patrick* et al *4978*, Polk Co., TN

C. major Walt.	2567	KM347940	*Schilling 07-DNA2567,* Unicoi Co., TN
C. major Walt.	3837	KM347941	*Schilling 2013-02*, Campbell Co., TN
C. pubescens Ell.	3095	KM347942	*Estes 9327*, Morgan Co., TN
C. pubescens Ell.	3830	KM347943	*Estes* et al *9326*, Cumberland Co., TN
C. rosea Nutt.	3838	KM347944	UT Garden 6/21/2013, Garden plant
C. tinctoria Nutt.	3831	KM347945	*Estes 4897*, Giles Co., TN
C. tinctoria Nutt.	3839	KM347946	UT Garden 6/21/2013, Garden plant
C. tripteris L.	3525	KM347947	*Floden 1719*, Campbell Co., TN
C. tripteris L.	3832	KM347948	*Estes and Beck 8481*, Bledsoe Co., TN
COSMOS Blume			
C. bipinnatus Cav.	3851	KM347948	*Schilling 2013-04*, Knox Co., TN
C. sulphureus Cav.	3870	KM347949	*Schilling 2013-06*, Knox Co., TN

Results and discussion

The newly obtained ITS sequences for Coreopsideae ranged in length from 631-640 bp. Sequences of *Coreopsis* were 631-640 bp in length; of *Bidens* were 635-639 bp, and of *Cosmos* were 638-639 bp. Many of the sequences from samples of *Coreopsis*, and some of *Bidens*, had polymorphisms. The ITS sequences for a subset of sect. *Coreopsis* species (*C. grandiflora*, all but one sample of *C. lanceolata*, and *C. pubescens*) all showed evidence of a length polymorphism involving a poly-G region at position 498-503 in the aligned matrix; the inferred lengths of the poly-G region for these samples were 5 and 6 bp. All other species of *Coreopsis* had a poly-G of 4 or 5 bp in length for this region, except for the sample representing a possible new taxon from limestone glades in Florida that had a poly-G region of 6 bp in length at this position. There were also length polymorphisms at different positions in the samples of *C. auriculata*, *C. major*, and *C. tinctoria.* Many samples of *Coreopsis* also had numerous positional polymorphisms (e.g., detected by the presence of a double peak; also called SNAPs, Whittall et al. 2000). The ITS sequences for some samples of *Bidens* also showed evidence of polymorphisms. Sequences of all samples of *B. discoidea* and *B. frondosa* shared a common suite of numerous positional polymorphisms and a length polymorphism. In contrast, samples of *B. aristosa* and *B. polylepis* had ITS sequences that were identical to one another and contained only three positional polymorphisms and no length polymorphisms. Sequences of the samples of *B. bipinnata*, *B. cernua*, and *B. tripartita* almost completely lacked any polymorphisms.

Many of the species of Coreopsideae from Tennessee were not well separated from close congeners based on ITS sequence results (Fig. 1). Within *Coreopsis*, there was little differentiation among numerous samples of sect. *Coreopsis*, including *C. auriculata*, *C. basalis*, *C. grandiflora*, *C. lanceolata* (both var. *lanceolata* and var. *villosa*), and *C. pubescens*; any differences involved whether or not there was a polymorphism at a given position. The separation of some samples of the section drawn from GenBank records (*C. auriculata* AY553677/ AY553678; *C. basalis* AY553705/ AY553706; *C. grandiflora* AY553707/ AY553708) seem likely to reflect errors in sequence reads rather than real differences, and these are relatively old records dating from when ITS1 and ITS2 were sequenced individually. An undetected length polymorphism could easily give rise to inaccurate base calls in the downstream sequence. A GenBank sample (GU724273) labeled *C. tinctoria* placed within this clade seems likely to be a misidentification. There was also little differentiation of our samples of *C. delphiniifolia* and *C. major*, although the GenBank samples labeled for each of these species were different (it appears that the GenBank sample for *C. major* might be misidentified and actually represent a *C. tripteris*). *Coreopsis delphiniifolia* has been proposed to be an allopolyploid involving *C. major*, *C. tripteris*, and *C. verticillata* L. (Smith 1976), but there was no evidence in the sequence or pattern of polymorphisms of either of the latter two species. A series of clones of a sample of *C. verticillata* deposited at GenBank (labeled there as *Coreopsis* sp. 1154; identified by the collector as *C. verticillata*, M. Vincent pers. comm.) showed some variability, but the variants appear

to represent autapomorphic changes. Within *Bidens*, there was greater differentiation among most but not all species. The exceptions involved *B. aristosa* and *B. polylepis*, which had identical sequences, and *B. discoidea* and *B. frondosa*, for which the ITS sequences were essentially identical. Both of these appear to be closely related (or perhaps not completely differentiated) pairs of species. For Coreopsideae and particularly for *Coreopsis*, these results suggest that although the ITS marker can narrow the possibilities for species level identification, it does not provide a suitable marker to distinguish uniquely several of the species.

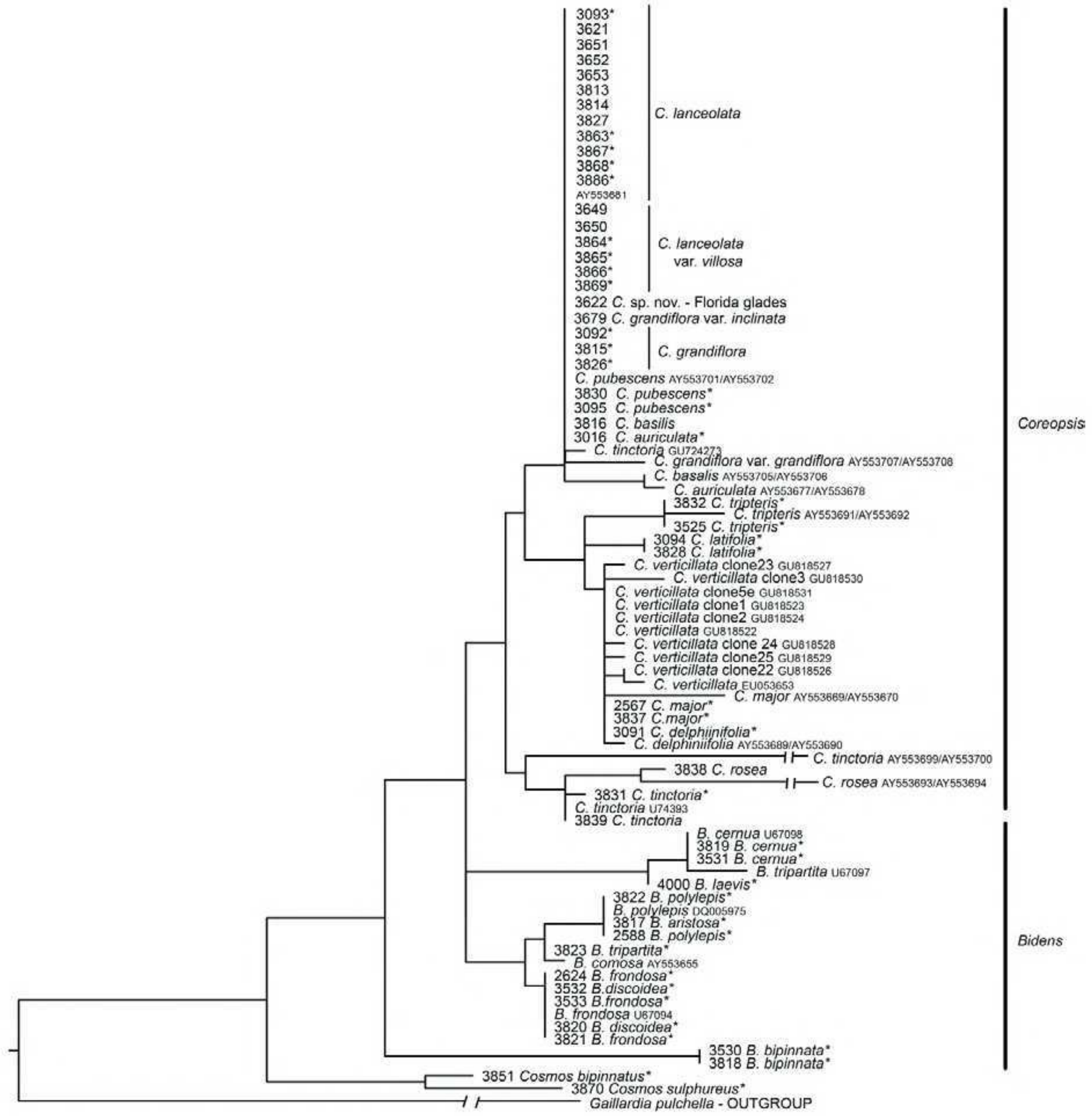

Figure 1. Maximum likelihood bootstrap tree (500 replicates) showing relationships of species of Coreopsideae based on ITS sequence data, using *Gaillardia pulchella* (Helenieae) as the outgroup. Newly obtained sequences designated by DNA number preceding species name (Table 1); GenBank numbers for other sequences follow species name. * = sample from Tennessee.

The low level of differentiation for the ITS marker at the species level within *Coreopsis* sect. *Coreopsis* complicates assessment of the potential distinctiveness of material from limestone glade habitats in Florida that is distinct morphologically from *Coreopsis lanceolata* (Johnson et al. 2013). There was, however, one clear difference between the Florida glade sample and the samples of *C. lanceolata* in the lack of any evidence in the former of a length polymorphism for the poly-G region that is almost universally present in other samples of not only *C. lanceolata* but also of *C. grandiflora* and *C. pubescens*. In addition, the Florida glade sample also exhibited a positional polymorphism at position 231 (A/G=R) that was instead present as an unambiguous G in all samples of *C. lanceolata*. By contrast, the sample of *C. grandiflora* var. *inclinata* from the Ketona Glades in Bibb Co., Alabama, had an ITS sequence that was identical to that of other samples of *C. grandiflora*. Thus even the small differences in ITS sequence in the Florida limestone glade sample appear to characterize it as distinctive, given the overall uniformity of the group for this marker.

The potential usefulness, but also the limitations, of utilizing the ITS region as part of a universal barcode can be seen in the results of this study. The ITS region remains popular for molecular systematics studies (e.g. Wang et al. 2014), in part because it is easily obtained even from herbarium material, as was the case here. All of the herbarium specimens that were extracted provided results, including one almost 40 years old. The ITS sequence data would allow near, if not exact, placement of all of the samples. However, the length polymorphisms, which were present in many of the samples of *Coreopsis* and also in some of *Bidens*, would complicate easy interpretation of results from direct sequencing. The database represented by GenBank samples continues to be problematic, both because of apparent misidentifications of material and also because of low quality of some older sequences. Nevertheless, as a quick and straightforward approach, surveying the ITS region as part of an initial analysis of species diversity appears to be a useful tool.

ACKNOWLEDGEMENTS

Thanks to J. Schilling for laboratory assistance, J. May and V. Brown of the University of Tennessee MBRF for help with DNA sequencing, M. Vincent for providing information on the GenBank record of *Coreopsis verticillata*, A. Johnson of the Florida Natural Areas Inventory program for providing material of the limestone glade populations related to *Coreopsis lanceolata*, and the numerous collectors who have helped to assemble a broad and useful herbarium at TENN. Financial support provided by the Hesler Fund of the University of Tennessee Herbarium.

LITERATURE CITED

Anderberg, A.A., B.G. Baldwin, R.G. Bayer, J. Breitwieser, C. Jeffrey, M.O. Dillon, P. Eldenäs, V. Funk, N. Garcia-Jacas, D.J.N. Hind, P.O. Karis, H.W. Lack, G. Nesom, B. Nordenstam, Ch. Oberprieler, J.L. Panero, C. Puttock, H. Robinson, T.F. Stuessy, A. Susanna, E. Urtubey, R. Vogt, J. Ward, and L.E. Watson. 2007. Compositae. Pp. 61–588, in J.W. Kadereit and C. Jeffrey (eds.). The Families and Genera of Vascular Plants. VIII. Flowering Plants, Eudicots, Asterales. Springer, Berlin.

Chester, E.W., B.E. Wofford, D. Estes, and C.J. Bailey, Jr. 2009. A Fifth Checklist of Tennessee Vascular Plants. Sida Bot. Misc. No. 31. Botanical Research Institute of Texas, Fort Worth.

Crabtree, T. 2012. Tennessee Natural Heritage Program rare plant list. <http://www.tn.gov/environment/na/pdf/plant_list.pdf> Consulted 2 Feb 2014

Crawford, D., M. Tadesse, M.E. Mort, R.T. Kimball, and C.P. Randle. 2009. Coreopsideae. Pp. 713–730, in V.A. Funk, A. Susanna, T.F. Stuessy, and R.J. Bayer (eds.). Systematics, Evolution, and Biogeography of the Compositae. International Association for Plant Taxonomy, Vienna.

Johnson, A.F., W.W. Baker, L.C. Anderson, and A.K. Gholson, Jr. 2013. Flora of calcareous upland glades in Gadsden and Jackson counties, Florida. J. Bot. Res. Inst. Texas 7: 475–494.

Schilling, E.E. 2013. Barcoding the Asteraceae of Tennessee, tribes Cardueae and Vernonieae. Phytoneuron 2012-99: 1–6.

Schilling, E.E. and A. Floden. 2012. Barcoding the Asteraceae of Tennessee, tribes Gnaphalieae and Inuleae. Phytoneuron 2012-99: 1–6.

Schilling, E.E., R.J. LeBlond, B.A. Sorrie, and A.S. Weakley. 2007. Relationships of the New England boneset, *Eupatorium novae-angliae* (Asteraceae). Rhodora 109: 145–160.

Smith, E.B. 1976. A biosystematics survey of *Coreopsis* in eastern United States and Canada. Sida 6: 123–215.

Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. 30: 2725–2729.

Wang, X-C., C. Liu, L. Huang, J. Bengtsson-Palme, H. Chen, J.-H. Zhang, D. Cai, and J-Q Li. 2014. ITS1: A DNA barcode better than ITS2 in eukaryotes? Mol. Ecol. Res. DOI: 10.1111/1755–0998.12325.

Whittall J., A. Liston, S. Gisler, and R.J. Meinke. 2000. Detecting nucleotide additivity from direct sequences is a SNAP: An example from *Sidalcea* (Malvaceae). Plant Biol. 2: 211–217.